The Ruth H. Hooker Research Library

and Technical Information Center

A CAMPUS-WIDE INFORMATION UTILITY AT THE NAVAL RESEARCH LABORATORY

Laurie E. Stackpole and Roderick D. Atkinson
Ruth H. Hooker Research Library and Technical Information Center
Naval Research Laboratory
Washington, D.C. 20375-5334

- **INTRODUCTION**

  The Ruth H. Hooker Research Library and Technical Information Center meets the information needs of Naval Research Laboratory's (NRL) research community, consisting of about 3,500 Federal staff and about 1,200 contractors. Since 1983, end users have been able to search the Library's online catalog both on site and remotely over the campus-wide network, known as NICENET. However, all other online searching continued to be performed by library reference staff, primarily using DIALOG, STN and DTIC. This situation began to change in 1988 when the Library introduced CD-ROM databases for end-user searching in both its reference area and its Microcomputer Software Support Center. Users responded favorably to the CD-ROM products, enjoying the freedom to explore that comes with performing their own searches.[1]

  In 1990, the Library conducted a user needs analysis to develop specifications for a replacement automated library system. This analysis was instrumental in establishing support for the development of a campus-wide information system for in-office access to library-based CD- ROMs and other in-library and external resources. As a result of interviews with 46 individuals representing a cross section of research interests, the Library was able to demonstrate that scientists throughout the laboratory wanted access to information resources from their own computers and workstations. Several key findings emerged, namely, that the Library should implement a system to:
    - Provide subject and author access to journal articles as well as books;
    - Allow users to request materials as part of an online search;
    - Offer access to multiple databases, both bibliographic and informational;
    - Store full text files, such as journal articles or handbooks, for downloading;
    - Provide access to the catalogs of other libraries and to external databases.

  While proceeding with plans to meet some of these objectives by identifying and procuring a state-of-the-art library system, the Library began to look at other approaches that could complement the capabilities of existing commercial library systems. Added support for these efforts came from the involvement of the Library in the Working Group tasked with planning for a fiber optic upgrade to the existing campus network backbone. The Working Group endorsed the concept of a library-provided information utility to:
    - Provide researchers with access to local computer systems;
    - Act as a gateway to remote systems;
    - Integrate these functions with library materials and services;
    - Make all this information available to researchers at their computers or workstations.

- **DESIRED FUNCTIONALITY OF AN INFORMATION UTILITY**

  To demonstrate the benefits of such an information utility, it may help to describe how a scientist might use electronic information. The scenario underlying the information utility concept, which at NRL is named InfoNet, envisions a scientist working at her computer writing a research paper. She gets to a point where she needs more information about the work of a colleague. She sends an E-Mail message to one of her peers and

gets an almost immediate response, citing two published papers mentioning Dr. X whose work is also of potential interest. The scientist is able to immediately query the InfoNet to find out if the cited articles are available. One is in the library collection in electronic format and is available for immediate downloading; the other is in the collection and will be faxed to her computer terminal within hours.

The scientist then decides to search one of several bibliographic databases available through InfoNet for papers by Dr. X. She finds only one and decides to learn more about Dr. X by searching a biographical database from a commercial information utility linked to the InfoNet, e.g., DIALOG. She finds Dr. X is affiliated with a major research institution. While she is on DIALOG, she also performs a citation search on Dr. X's paper and finds it has been cited by a colleague whose work she greatly respects. Satisfied that Dr. X is a highly regarded professional, she returns to the Library's online catalog through the InfoNet, where she finds that the journal in which Dr. X has most recently published is not one to which the Library subscribes. However, she sees on the InfoNet that individual articles reside on a remote server at a university accessible through InfoNet and are available for viewing.

It turns out that Dr. X has written about a related area, one which the NRL scientist is not terribly familiar. So at this point after printing out the Dr. X's article at her printer, she decides she would like to find out if there is anyone at NRL who can bring her up to speed. She performs a subject search on a database of NRL publications and searches the CD-ROM database of Work Unit Summaries. She finds two fellow NRL scientists who she thinks could help her. She consults the online personnel locator for their E-Mail addresses, sends them messages, and returns to work on her paper. All this took considerably less than one hour and was accomplished without her leaving her computer. The information utility supported all her information-seeking activities with a menu-driven user interface and pre-programmed access to remote systems.

In 1990, when this concept was being developed, it was not easy to see how the Library would make much of this happen. While access to the existing or replacement library system was straight forward, end user access to other Laboratory databases and to remote computers as a part of a suite of library services did not appear easy. There seemed to be only three possibilities: the Lawrence Livermore model, with dedicated gateway software running on a local computer[2]; the DTIC Search Maestro and CompuServe model, with gateway software running on a remote computer; and the library system model as exemplified by the STILAS (Scientific and Technical Information Library Automation System) Remote Interface Module.

However, several diverse developments were taking place in the early 1990's. Taken together, three of these provided the Library with a framework to move aggressively toward the information utility concept.

- **ADVANCES IN NETWORKING INFORMATION**

  One encouraging development was the in-library networking of CD-ROM databases at installations such as the Nimitz Library at the U.S. Naval Academy[3]. This work showed it was possible to provide menu access to multiple CD-ROM products and allow multiple users to simultaneously search a single CD-ROM. If it is possible to do this on an in-library LAN, would it not be possible to provide the same capability over a campus-wide network?

  Another encouraging development was the work going on at many universities to implement Campus-Wide Information Systems (CWIS). Institutions such as MIT[4], RPI[5], Dartmouth[6], Carnegie Mellon[7], and Case Western Reserve[8] were providing students and faculty with access to a wide variety of informational material and databases over their campus networks. Offerings were wide ranging and varied from university to university. Typically the CWIS offered access to course descriptions, class schedules, event calendars, campus publications, and often one or more bibliographic databases. In many cases, the CWIS was operated by the computer center and developed independently of the Library. In other cases, such as the RPI InfoTrax and the Carnegie Mellon Mercury Project, it was being developed as an extension of existing library resources, showing that it was feasible to conceive of a library-operated information utility.

  The final enabling development was the rapid expansion of the Internet. Suddenly it seem everyone was on the Internet. Not only was the NRL NICENET tied into the Internet, but information providers, such as OCLC and DIALOG, had an Internet address. The sell-out sessions at meetings such as SLA and ALA were

the sessions on using the Internet; journal articles about what was on the Internet abounded[9]. Not only was there more information available than most people had ever conceived of, now it was possible to quickly and easily get to it.

THE INFONET

- **ENVIRONMENT**

  To accomplish the goal of making a local information utility for use at NRL, the Library and its contractor, Kestrel Associates Incorporated, had to deal with three basic issues. The first of these was the MS-DOS operating system adopted by most commercial CD-ROM publishers. The second was the wide variety of equipment in use at NRL, including: hundreds of dumb terminals that provide access to the Laboratory's mainframe computers, a SMTP and DECNet E-Mail system and administrative files; a few thousand PC compatible and Macintosh computers, in heavy use in both research and administrative areas; and a large number of VAX, SUN and other UNIX workstations used by researchers. The third was the need to provide gateways to remote information utilities, both at NRL and throughout the world, running on a variety of systems.

- **IMPLEMENTATION PLAN**

  From a management perspective, a phased implementation to the InfoNet design appeared the most promising. This modular sort of approach allowed functions to be brought online and thoroughly tested before being integrated with the system as a whole. As a result, problems could be more easily isolated and identified and system-wide disruptions would be infrequent.

  A problem often associated with the phased approach to system implementation is that it can lead to a myopic view of the system, i.e., each module might be seen as an end unto itself. The final product may function more as a number of small systems, each accessed independently and providing its own services rather than achieving the final objective of a fully integrated system. In developing the InfoNet, this myopic syndrome was avoided. Each phase was viewed from both a small systems perspective and from the broader viewpoint of how it would integrate with the InfoNet system. "While a detailed knowledge of LAN technology is an asset, it is far more important that a LAN administrator have a clear conceptual overview...."[10] The ultimate goal for the InfoNet was clear: unify electronic information resources under a single menuing system, and make the system accessible to researchers in their offices and laboratories and independent of client computing platform.

  The InfoNet was brought online in three distinct phases[11]:
  Phase I:
  implement a library LAN with networked CD-ROM databases and office productivity software, e.g., word processing, local E-Mail[12], database management, etc. for PCs and Macintosh microcomputers;

  Phase II:
  integrate the library LAN with the campus-wide network to provide NRL with network access to CD-ROM databases independent of computing platform, provide staff with Internet E-Mail[13], and make campus-wide and Internet information resources and databases accessible from within the library under a single menu driven system;

  Phase III:
  provide the campus-wide network with a menu driven system which includes: MS-DOS based CD-ROMs, library information databases running under SUN UNIX (i.e., the library's Online Public Access Catalog and a NRL Authors' Citations Database), DEC VMS based Management Information Databases run by NRL's MIS Branch, and information resources and databases accessed on the Internet.
  The InfoNet can be accessed from anywhere on NRL's campus-wide network, regardless of computing

workstation or dumb terminal[14]. In brief, the InfoNet links to the FDDI/Ethernet campus backbone primarily using the TCP/IP suite of protocols. UNIX workstations and dumb terminals use native Telnet to access the InfoNet while computers which lack native TCP/IP support, e.g. MS-DOS users, use free Telnet software distributed by the library and modified for InfoNet usage[15]. Macintosh users take advantage of native AppleTalk protocols to access the InfoNet and are supplied with software distributed by the library for AppleTalk connectivity. The InfoNet consists entirely of off-the-shelf PC hardware and software, with a few minor software modifications.

To provide enhanced client access to Internet information resources and databases, the InfoNet makes extensive use of state-of-the-art "Knowbot" (Knowledge Robot) prototypes, such as Gopher[16], NNTP[17], WAIS[18], and Archie[19]. Knowbots are "programs that, once activated, wander through the [Internet] looking for information and [return] it to their electronic masters."[20] The Knowbot prototypes query unrelated host systems and display the information to the end-user in a standard and familiar format. By separating the client search software from the database host, end-users are no longer required to learn the interfaces of unrelated systems. (Unfortunately, CD-ROM publishers do not support Knowbots at this time.) These programs are continuously updated and help to tame the Internet. In so doing, they have made world-wide information systems increasingly accessible. Recent statistics on InfoNet usage have shown that Internet search applications now represent half of all searching performed on the InfoNet.

CD-ROM DATABASES

- **DATABASE LICENSING**

  CD-ROMs are almost always licensed to the customer and usually restrict the use of data far beyond that of copyright law. Since CD-ROMs are licensed to the customer and not sold, the publisher is at liberty to restrict the use of the product in the license, assuming there is agreement from the customer. Unfortunately the publishing industry lacks any kind of standard licensing agreement and individual publishers vary greatly in how they license their products. Generally speaking, however, licenses usually restrict the copying of data from the CD-ROM onto other media and usually require that the product be returned or destroyed after either a new disk is sent to the customer, or the subscription is canceled. Moreover, network licensing agreements often control the number of users, workstations, buildings, and/or sites connected to the network and restrict the use of modems for remote network access.

  Publishers seldom understand how networks disseminate information or what possibilities exist for controlling information access. For example, the majority of the NRL research community is spread out over approximately 150 low-rise buildings, interconnected by a campus network. The actual number of computer workstations attached to the campus network exceeds the total number of employees. Although over ninety percent of employees are at a single site in Washington, D.C., NRL has facilities and employees in other states. Further, many researchers continue their work at home in the evenings and on week-ends and holidays. As a result, the common conditions applied in CD-ROM licensing agreements are not viable in such an environment. To compensate for this, licensing terms must usually be negotiated with each publisher individually with network access to CD-ROM databases limited usually to no more than five simultaneous users controlled through metering software. In addition, the InfoNet restricts network access to IP (Internet Protocol) addresses used at NRL. The cost for a concurrent license varies, but it is seldom more than twice the price of a stand-alone version and permits around five simultaneous users. For many vendors, multi-site and dial-in access licensing has yet to be developed and the InfoNet will be breaking new ground as we move into the next phase of implementation.

  When discussing network licenses with CD-ROM vendors, it is important to emphasize the needs of the administrators and end-users and not to succumb to the wishes of the publishers should the two be inconsistent. It is in the best interest of both the library and the publisher that an equitable arrangement be negotiated. Few CD-ROM vendors understand either the possibilities or the limitations of networking CD-ROMs and therefore any network license agreement that does not conform to the library's needs should be negotiated. In almost every case CD-ROM publishers were receptive to the suggestions of the NRL library; licenses and pricing were altered to reflect the needs of the InfoNet.

- ## CD-ROM SEARCH ENGINES AND NETWORKING

  Most CD-ROM search engines are not well adapted for use on a campus-wide network because such networks commonly use the TCP/IP suite of protocols. These protocols follow standards which are often ignored the CD-ROM search software. The result is that end-users sometimes run into difficulties using the search software. Some of the problems which NRL has encountered include:
  - Most CD-ROM search software assume a stand-alone PC and send 25 lines of text. However, TCP/IP terminal emulation is limited to 24 lines, regardless of whether the end user is utilizing a PC or any other type of system; therefore, the commands the search software puts on the 25th line are invisible to the end user on a TCP/IP network.

  - Color codes are often used to indicate cursor position. Some search software generate highlighting via color code changes rather than standard inverse video control characters; the result is that end users cannot see changes in menu selections or even determine the position of the cursor.

  - The use of "alt" keys is also non-standard in TCP/IP networks. No TCP/IP networking packages support "alt" keys on PCs, and non-PC keyboards do not even possess "alt" keys. In addition, only Function keys "F1" through "F5" are fully supported by most TCP/IP networking packages. As a result, end users must learn complicated key combinations to execute even simple search commands.

  - The search software often restricts the end user to searching only one CD-ROM database at a time. While this is both acceptable and appropriate in a stand-alone environment, this is highly inefficient and cumbersome on a network.

  - Although most software conforms to both the ISO9660 and the Microsoft MSCDEX standards, a few search engines, like that of Online Software, Inc., circumvent the standards and communicate information directly to the CD-ROM drive. This precludes use on TCP/IP based networks.

  To compensate for search software dependence on "alt" keys and non-standard function keys, the library distributes software for PC-compatible and Macintosh microcomputers. MS-Kermit is free MS DOS based TCP/IP software from Columbia University and has been modified by the NRL Library to emulate a PC keyboards. For Macintosh users, ONLAN MAC is distributed as part of the license agreement for Novell Access Servers and provides full PC keyboard and monitor capabilities on a Macintosh connected to the InfoNet. In addition, UNIX and dumb terminal users are provided surrogate keys which have been remapped at the host to provide keys not found on their native keyboards.

  The use of these techniques on InfoNet make the CD-ROM search engines usable on all platforms. However, applying these techniques is not always intuitive for the use nor are the results fully satisfactory. The real solution to the problem is for CD-ROM publishers to conform to networking standards when designing CD-ROM search software.

- ## PLANS FOR THE INFONET

  The Library staff takes advantage of three primary tools for planning InfoNet development and improvements: usage statistics, end-user comments, and industry surveys. Statistics are recorded automatically on a per application basis every time any of the InfoNet's resources are accessed and serve as essential elements for focusing collection development efforts. Comments from InfoNet users about missing features and information are recorded for careful consideration. Library staff continuously monitors the networking and publishing community for new products and directions. From these sources it has been determined that the InfoNet needs to concentrate efforts on the following areas:
  - Dial-in and remote access to all the services of the InfoNet both from home and from NRL sites in Monterey, Orlando and Mississippi;
  - Support for a X-Windows interface;
  - "Electronic stacks" of journal images in full-page format[21];
  - Additional sophisticated "Knowbots" to search the Internet.

- ## CONCLUSION

Recent advances in technology have made it possible to bring information to the researcher where he or she needs it most: at the desktop and on the researcher's computing platform of choice. For NRL, the InfoNet serves as a networked information utility which unites information resources from a variety of computing hosts under a single menuing system, and which itself is accessible to all NRL computing platforms. The InfoNet itself is easily expanded and consists entirely of off-the-shelf PC and networking hardware and software. The InfoNet proves that CD-ROM databases can be integrated with campus-wide networks and made available to users independent of computing platform, although not all search features can be used by all varieties of computers. The InfoNet also proves that researchers have a need for facilitated access to the Internet and the InfoNet can meet those needs using state-of-the-art information-seeking software.

- **NOTES**

1. Laurie E. Stackpole, CD-ROM in a Federal Scientific-Technical Library. *CD-ROM EndUser*, Vol. 2, No. 2, pp. 60-62, June 1990.

2. Hilary D. Burton, The Livermore Intelligent Gateway: An Integrated Information Processing Environment, *Information Processing & Management*, vol. 25, no. 5 pp. 509-514 (1989).

3. Ruth M. Hennessy, "Midshipman's Use of CD-ROM at the Naval Academy... in *CD-ROM: Facilitating Electronic Publishing* by Linda W. Helgerson (Van Nostrand Reinhold, 1992), pp. 301-302.

4. Timothy J. McGovern, "TechInfo -- Public Information at MIT" in *Using Computer Networks On Campus*, Papers from the First Annual Conference 1990, edited by Les Lloyd (Mecker, Westport, CT, 1991).

5. Barbara Lockett, "Rensselaer Libraries InfoTrax System" in *Using Computer Networks on Campus*, Papers from the First Annual Conference 1990, edited by Les Lloyd (Meckler, Westport, CT, 1991).

6. Gregory A. Finnegan, Wiring Information to a College Campus: A Port for Every Pillow, *Online*, vol. 14, no. 2, pp. 33-40 (March 1990).

7. The Mercury Electronic Library and Library Information System II: The First Three Years, Carnegie Mellon University Mercury Technical Reports Series, no. 6, Carnegie Mellon University, Pittsburgh, PA (February 1992).

8. Ray Metz, Integrating Local Library Systems and Services into a Campus Network Environment, *Computers in Libraries*, vol 10, no. 6, pp. 18-20 (June 1990).

9. Example: Aggi W. Raeder and Karen L. Andrews, Searching Library Catalogs on the Internet: A Survey, *Database Searcher*, pp. 16-31 (September 1990).

10. Richard W. Boss, *Library Technology Reports*, vol. 28, no.3, p.340. (May-June 1992).

11. In designing the InfoNet, extensive use was made of software found on the Internet. "Anonymous FTP" sites, that is host sites on the Internet which support public downloading of software, are noted where appropriate. InfoNet administrators use NCSA FTP (ftp.ncsa.uiuc.edu) to download up-to-date software.

12. Pegasus E-Mail for Novell v.2.34 for PCs and v.2.0 for Macintosh, Freeware, Dunedin, New Zealand. (Distributed by the University of Hawaii, 128.171.17.7).

13. Charon v.4.0, SMTP (Internet) E-Mail Gateway for Pegasus E-Mail, Freeware, Clarkson University. (128.153.4.2).

14. For readers interested in a technical overview of the hardware and software used on the InfoNet, refer

to: Laurie E. Stackpole, Roderick D. Atkinson and John Yokley, Campus-wide Network Access to CD-ROM Databases, Proceedings of the Thirteenth National Online Meeting, New York, 1992 (Learned Information, Inc., Medford, NJ).

15. MS-Kermit v.3.12, from Columbia University, (watson.cc.columbia.edu), modified by the library with scripting and keyboard mapping appropriate for MS-DOS based software applications.

16. Gopher II v. 1.05b, from the University of Minnesota. (boombox.micro.umn.edu).

17. Trumpet v. 1.05g, USENET (Internet) News Reader, from the University of Tasmania, Australia (tasman.cc.utas.edu.au).

18. WAIS, University of North Carolina, (samba.oit.unc.edu).

19. Archie, McGill University, Canada (archie.mcgill.ca).

20. J. T. Johnson, NREN: Turning the Clock Ahead on Tomorrow's Networks, *Data Communications*, vol.21, no.12, p. 58 (September 1992).

21. G. A. Story, L. O'Gorman, D. Fox, L. L. Schaper and H. V. Jagadish, The RightPages Image-Based Electronic Library for Alerting and Browsing, *Computer*, vol. 25, no. 9, (September 1992).